

Chapter 2. Methods

This chapter briefly summarizes some key issues related to the organization of the quality report and methods used to construct the measure set, conduct relevant data analysis, and report conclusions.

Selection of Measures and Data Sources

Formal input was received through an Interagency Workgroup and a presentation to the HHS Data Council. Representatives from a range of HHS agencies provided ongoing input throughout the development process of the first quality report through this temporary work group. Future reports may rely on a similar group to provide input as the measure set evolves. The NHQR Interagency Group includes representation from:

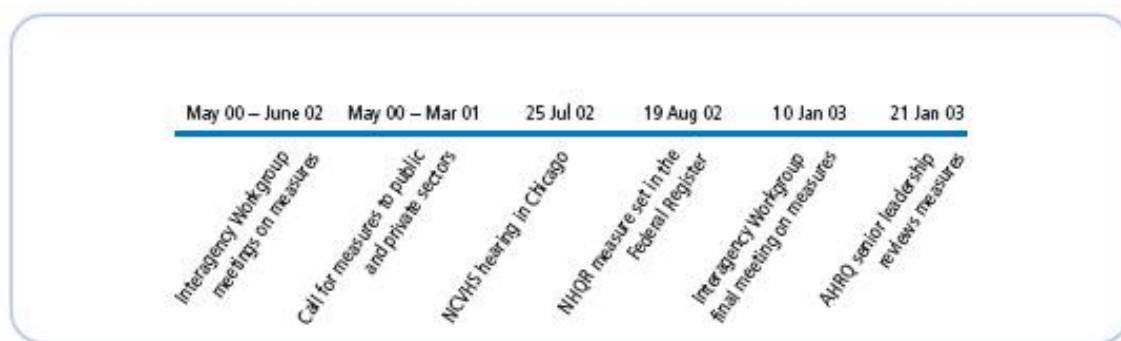
- Assistant Secretary for Planning and Evaluation (ASPE)
- Centers for Disease Control and Prevention (CDC)
- CDC-National Center for Health Statistics (NCHS)
- Centers for Medicare & Medicaid Services (CMS)
- Food and Drug Administration (FDA)
- Health Resources and Services Administration (HRSA)
- Indian Health Service (IHS)
- National Institutes of Health (NIH)
- Substance Abuse and Mental Health Services Administration (SAMHSA)

In order to select measures for the report, a subgroup was drawn from the above group—the NHQR Measures Workgroup—which included representatives from AHRQ, the National Center for Health Statistics, the Centers for Medicare & Medicaid Services, the National Institutes of Health, and the Office of the Assistant Secretary for Planning and Evaluation. This workgroup developed a “call for measures” that was sent to all relevant Federal agencies. The Institute of Medicine issued a complementary call for measures to the private sector. Those submitting measures also had to submit the name of a proposed data set. More than 600 measures were submitted for consideration in response to these calls (see Figure 1, on the next page, for abbreviated timeline for selecting measures).

The Measures Workgroup mapped the candidate measures into the fleshed-out conceptual framework. The measures within each category of care were evaluated for inclusion in two parts:

1. Measures were selected to maintain consistency with existing consensus-based measure sets where possible. For example, approximately 30 measures were submitted relevant to the management of diabetes. The National Diabetes Quality Improvement Alliance—a collaboration of the American Medical Association, the Joint Commission on Accreditation of Healthcare Organizations, and the National Committee for Quality Assurance—recently announced a set of eight recommended performance measures for the management of diabetes. The Interagency Workgroup focused on these eight measures for use in tracking the effectiveness of diabetes management for the report.

Figure 1. Timeline of the measures selection process



2. The workgroup assessed candidate measures using the following criteria:¹
 - **Importance.** What is the impact on health associated with the health problem assessed by the measure? Are policymakers and consumers concerned about this area of health care quality? Can the health care system meaningfully address this aspect or problem?
 - **Scientific soundness.** Does the measure actually reflect what it is intended to measure? Does the measure provide stable results across various populations and circumstances? Is there scientific evidence available to support the measure?
 - **Feasibility.** Is the measure in use? Can information needed for the measure be collected in the scale and time-frame required? How much will it cost to collect the data needed for the measure? Can the measure be used to compare different population groups?

Based on these assessments the workgroup identified a preliminary set of about 140 measures for the first quality report. This review included the following steps:

- Measures were sorted into the conceptual framework.
- Using the selection criteria above, measures not suitable were eliminated from the measure set.

- Agencies were given a complete list of the measures and asked to rate their measure submissions and those of others using a structured rating form containing the criteria.
- Agencies were then asked to submit any additional measures.

A particular effort was made to include both process measures that assess what happens to patients during their care and outcome measures that track what ultimately happens as a result of that care. Process measures are more direct assessments of the quality of particular care received and have been shown to be more sensitive for detecting differences between individual health care institutions.² However, adequate process measures with national data sources have not as yet been developed in many clinical quality areas. Outcome measures of quality have inherent methodological issues when used to judge quality.³ The NHQR Interagency Workgroup worked to select process measures that are closely linked to outcomes and outcome measures that are understandable, valid, and can, when appropriate, be adjusted for other factors such as severity of illness or age.

Following the workgroup's work on the measures, the preliminary measure set was reviewed by internal experts and senior management at AHRQ. It was then presented to reviewers within HHS, including the Quality Interagency Coordination Task Force (QuIC) and the HHS Data Council.¹ External feedback was obtained through two primary vehicles. The first was a hearing sponsored by the National Committee on Vital and Health Statistics in Chicago on July 25, 2002. The second was a call for feedback on the NHQR preliminary measure set in the *Federal Register* published August 19, 2002. This feedback was synthesized and reviewed by the NHQR Measures Workgroup with the goal of generating necessary additions, deletions, and alterations to measures in the measure set. This feedback was then reviewed by the NHQR Measures Workgroup and sorted for action. Action items were then forwarded on to the full NHQR Interagency Workgroup which met on January 10, 2003. The full group made a set of recommendations on the proposed final measure set to AHRQ senior leadership, which reviewed these recommendations on January 21, 2003. Following this review, the measure set for the first NHQR was updated and finalized.

The process for selecting home health measures differed from that for the other measures. The preliminary measure set (dated August 19, 2002) did not propose any home health measures. This was because AHRQ was working together with CMS to determine an appropriate set of measures for the CMS public reporting initiative on home health as well as this report. AHRQ and CMS decided that, in the short term, the Outcome and Assessment Information Set (OASIS) measures would be used as the initial measure set because there is more standardization around these measures than any other in home health care.

¹ The HHS Data Council coordinates all health and non-health data collection and analysis activities of the Department of Health and Human Services, including an integrated health data collection strategy, coordination of health data standards and health information and privacy policy activities. The HHS Data Council consists of all Assistant Secretary and Agency Administrator level HHS officials who have a direct reporting relationship to the Secretary, the HHS Privacy Advocate, and the Secretary's Senior Advisor on Health Statistics. It is co-chaired by ASPE and a rotating Operating Division (OpDiv) head; AHRQ is the current OpDiv co-chair. (See <http://aspe.hhs.gov/datacncl/> for more information.)

AHRQ convened the Home Health Quality Measures Technical Expert Panel (TEP) to review the set of OASIS home health quality measures as candidates for both the NHQR and the CMS home health care public reporting initiative. Accordingly, AHRQ convened a TEP October 21-22, 2002, to address these two independent but overlapping efforts being planned by CMS and AHRQ.

Based on the Home Health TEP input, including: the individual panelist prioritization lists (i.e., a significant proportion of panelists listed particular measures as priority items for inclusion), their written comments and the meeting discussion, and AHRQ's proposed 12 OASIS measures for reporting on the quality of home health care in the report, a Federal Register notice was published March 24, 2003 requesting public comments on these measures. Written comments were due by April 23, 2003.ⁱⁱ

Data Sources

This report is intended to track quality for the Nation over time. As such, it must rely on readily available, reliable and valid, regularly and consistently collected data at both the national and State levels. These requirements restricted the data sources that could be used for the report. When the call for measures was made, there was also an accompanying request for data sources for the proposed measures. During the developmental phase of the project, the workgroup devised a two-tiered scheme for characterizing possible data sources for the report. Each potential data source was classified according to the criteria presented in Table 1.

Table 1. Two-tiered categorization scheme for examining data sources

Tier I: Substantively relevant and nationally representative:	Tier II: Substantively relevant but:
<ul style="list-style-type: none"> • For the target population under consideration. • For a given population such as civilian, resident, noninstitutionalized, nursing home residents, etc. • And accurate and reliable with specified relative error. • With the capacity for multiple levels of detail. • With acceptable response rates. 	<ul style="list-style-type: none"> • Adjusted to compensate for limitations in national representation. • Data representative at the subnational level (such as State or Metropolitan Statistical Area). • Data not nationally representative but substantively important.

ⁱⁱ A Measures Background Appendix providing detailed information and rationale for inclusion of measures in the measure set is currently in development and will be available online at a later date.

This system of categorization helped to identify established, national data sources that are the standard for providing national estimates over time for the report. The data from these data sources provide estimates for the U.S. civilian, noninstitutionalized population.

Table 2 presents a list of these data sources and compares them to data sources presented in other national quality reporting efforts. This table illustrates how this report lines up with other established national efforts at assessing health care and health care quality. All of the data sources used for these national reports are also used in the NHQR. More detail on the data sources, populations, and other relevant information for the measures is contained in the Measure Specifications Appendix for this report.

Table 2. Comparison of data sources for the NHQR and other quality reporting efforts

NHQR Data Sources	Data Sources for Other Quality Reporting Efforts
<ul style="list-style-type: none"> Behavioral Risk Factor Surveillance System (BRFSS) Dialysis Facility Compare (DFC) Healthcare Cost and Utilization Project (HCUP) Health Plan Employer Data and Information Set (HEDIS®) HIV/AIDS Surveillance System Medical Expenditure Panel Survey (MEPS) Medicare Quality Monitoring System (MQMS) Minimum Data Set (MDS) National Ambulatory Medical Care Survey (NAMCS) National CAHPS® Benchmarking Database (NCBD) National Health and Nutrition Examination Survey (NHANES) National Health Interview Survey (NHIS) National Home and Hospice Care Survey (NHHCS) National Hospital Ambulatory Medical Care Survey (NHAMCS) National Hospital Discharge Survey (NHDS) National Immunization Survey (NIS) National Nosocomial Infections Surveillance (NNIS) National Nursing Home Survey (NNHS) National TB Surveillance System (NTBSS) National Vital Statistics System —and Infant Death Data (NVSS-I) National Vital Statistics System, Mortality (NVSS-M) Outcome and Assessment Information Set (OASIS) Quality Improvement Organization (QIO) Surveillance, Epidemiology, and End 	<p><i>Healthy People 2010</i></p> <ul style="list-style-type: none"> Behavioral Risk Factor Surveillance System (BRFSS) Continuing Survey of Food Intake by Individuals (CSFII), 1994-96 HIV-AIDS Case Surveillance System Medical Expenditure Panel Survey (MEPS) Monitoring the Future Study (MTF) National Ambulatory Medical Care Survey (NAMCS) National Crime Victimization Survey (NCVS) National Health and Nutrition Examination Survey (NHANES) National Health Interview Survey (NHIS) National Hospital Ambulatory Medical Care Survey (NHAMCS) National Hospital Discharge Survey (NHDS) National Household Survey on Drug Abuse (NHSDA) National Notifiable Disease Surveillance System (NNDSS) National Electronic Telecommunications System for Surveillance (NETSS) National Profile of Local Health Departments (NPLHD) National Survey of Family Growth (NSFG) National Vital Statistics System, Mortality (NVSS-M) National Vital Statistics System, Natality (NVSS-N) 1999 National Worksite Health Promotion Survey (NWHPS) School Health Policies and Programs Study (SHPPS) State Tobacco Activities Tracking and Evaluation System (STATE) STD Surveillance System (STDSS) United States Renal Data System (USRDS) Youth Risk Behavior Surveillance System (YRBSS) <p><i>National Committee for Quality Assurance</i></p> <ul style="list-style-type: none"> Health Plan Employer Data and Information Set (HEDIS®)

<ul style="list-style-type: none"> Results Program (SEER) United States Renal Data System (USRDS) 	<p><i>Commonwealth Report on Quality of Health Care in America</i></p> <ul style="list-style-type: none"> Behavioral Risk Factor Surveillance System (BRFSS) Commonwealth Fund International Health Policy Survey Community Tracking Study Household Survey (Center for Studying Health System Change) Health Plan Employer Data and Information Set (HEDIS®) Medicare Cooperative Cardiovascular Project Minimum Data Set (MDS) National Health Interview Survey (NHIS) National Health Interview Survey on Disability National Hospital Ambulatory Medical Care Survey (NHAMCS) National Immunization Survey (NIS) National Registry of Myocardial Infarction, American College of Cardiology National Survey on Nursing Homes (NewsHour with Jim Lehrer/Kaiser Family Foundation/Harvard School of Public Health) On-line Survey, Certification and Reporting (OSCAR) system; CMS Picker surveys of 272 self-selected hospitals Quality Improvement Organization (QIO) Surveillance, Epidemiology, and End Results Program (SEER) Vital Statistics
---	---

Significance Testing and Assessing Relevant Differences

With a range of conditions and measures, AHRQ established a systematic process for reviewing data and assessing relevant differences. Reported comparisons are for statistically significant differences unless otherwise noted. This process can be summarized as follows:

- All tables were generated for all measures with the appropriate estimates and standard errors as needed for the table analysis. Standard errors are numeric representations of the error that occurs because the estimate is based on a sample of a larger population.
- Statistical testing was conducted on the estimates. The tests done were two-tailed t-tests of significance at the alpha level of 0.05. All of the data that are highlighted in this report meet this statistical criterion. The testing included these steps.
 - For national tables, differences between estimates for subgroups and the identified comparison (reference) group were tested for statistical significance.
 - For national tables with data over time, the least recent year was used as the reference and subsequent years were tested versus that reference year.
 - For State tables, States were compared with the national average. (Please note that these differences between States and the national average were computed solely

to highlight opportunities for improvement nationally rather than as assessments of the performance of individual States.)

- The report team systematically assessed the estimates for changes over time nationally, between relevant subgroup populations nationally and between States and the national average. Appropriate differences were highlighted for review.
- The highlighted differences noted by AHRQ's clinical team were then reviewed by AHRQ senior leadership and a team of national experts in areas of the NHQR framework.

The report cites various quality improvement studies. On such studies, none of the quality indicator information discussed is part of the official report dataset. Therefore, AHRQ did not perform statistical testing on these indicators.

Data Suppression

Sometimes not all the data collected from surveys, medical records, or administrative sources can be presented. Oftentimes, this is due to the small number of cases in particular categories of reporting that are not considered reliable. Even more important, presentation of these numbers may jeopardize confidentiality. When data are collected and analyzed but not presented for reasons such as these, it is called suppression. Different data collection systems apply different criteria to suppress data that are deemed unreliable ranging from no suppression of data to suppressing data through complicated algorithms. The rule employed for data suppression for this report was to adhere to the rules specified by the data source from which the measure was derived. (Detailed information on each of the data sources is contained in the Measure Specifications Appendix.) For most data sources, there were two main data suppression criteria: 1) cell values based on unweighted N less than 30 and 2) relative standard errors greater than 30% when appropriate. Some data systems have more stringent suppression criteria, and their criteria were maintained in this report for their data. For example, the National Health Interview Survey (NHIS) suppresses data with cell sizes less than 50, and the Medical Expenditure Panel Survey (MEPS) suppresses data with cell sizes less than 100. NHIS and MEPS data presented in this report adhere to these more conservative standards.

A general exception to these data suppression criteria is for data that encompass population counts such as vital statistics. Such data typically have their own suppression criteria. For example, mortality statistics based on fewer than 20 deaths are typically suppressed; data presented in this report adhere to this criterion.

Limitations of the Measure Set and Analysis

This report explicitly relies on measures that have broad national consensus. It is built on measures focused on the Nation's health care priorities as determined by associated morbidity and mortality as well as the opportunity for improvement. In order to fulfill the congressional

mandate, the report is also based on measures for which national data are available. Consequently, in many areas covered in the report, it does not rely on the latest research findings on quality of care measurement. Wherever possible, we have attempted to cite appropriate and promising research on quality measurement.

The legislative charge of the report is to provide a “snapshot” of the Nation’s quality of care. As a snapshot, it is intended to neither test hypotheses nor prove causation for the apparent gaps in quality but to provide a benchmark for future analyses. The statistical analysis performed for the report is in keeping with this goal and the scope of the report.

Bivariate analyses were performed for the report as the first step to presenting where the Nation stands in terms of quality of care. While multivariate analyses are preferable for more detailed exploration of the data, they were not feasible in the timeframe of the initial report. Future editions of this report will include analyses of the relationship between the multiple factors that may explain differences between and among these populations. Adjusting for known contributing factors, such as age and sex, can help present more accurate results. In addition, nearly half of the measures in the report (67 of 147) present adjusted data. The measure set does not include any measures of structural quality.

Also, this is not a report on the level of quality at individual hospitals or doctors’ offices. It is intended as a tool for Federal and State policymakers, and therefore, it tracks quality at the national and, wherever possible, State levels. Data are reported in the body of the report and in the detailed appendix tables at the national and, if data exist, State levels. This means that quality is discussed in terms of the following: a) national performance over time (if trend data are available); b) national performance for key priority subpopulations; and c) variation between States.

This is the first in a series of annual reports to Congress. Comments received during the design and public comment stages of the report development have helped improve the design of this first report. However, it will only be possible to implement recommendations in future reports. Potential areas for improvement are discussed in the report, and feedback on this first report is expected to help shape subsequent editions.

Additional Information

For information on the specifications for the measures and the data sources, readers are encouraged to consult the Measure Specifications Appendix.

Readers are also encouraged to consult the AHRQ Web site for more information on the report and its design at www.ahrq.gov.

References

¹Institute of Medicine. Envisioning the National Health Care Quality Report. Washington, DC: National Academies Press; 2001.

²Mant J, Hicks N. Detecting differences in quality of care: the sensitivity of measures of process and outcome in treating acute myocardial infarction. *BMJ* 1995;311(7008):793-6.

³Orchard C. Comparing healthcare outcomes. *BMJ* 1994;308(6942):1493-6.

